|  |
| --- |
| **ĐỒ ÁN CUỐI KỲ “AI CƠ BẢN VÀ ỨNG DỤNG”**  Tên đề tài: Nhận Diện Tiền Giấy Việt Nam  Giảng viên: ThS. Lương An Vinh  Họ và tên sinh viên: Trần Đoàn Thanh Sang  Mã sinh viên: DH51904363 |

**MỤC LỤC**

**Phần 1: Giới thiệu, mô tả đề tài**

1. Giới thiệu
2. Mô tả bài toán cần giải quyết
3. Dữ liệu và nguồn dữ liệu
4. Các đặc trưng sử dụng
5. Phương pháp đánh giá..............................................................
6. Cách bố trí thực nghiệm
7. Phương pháp học

**Phần 2: Ứng dụng**

1. Xác định yêu cầu bài toán
2. Các chức năng
3. Giao diện

**Phần 3: Kết luận**

1. Ưu điểm
2. Nhược điểm

**Phần 1: GIỚI THIỆU, MÔ TẢ ĐỀ TÀI**

1. **Giới thiệu**

* Với sự phát triển hết sức mạnh mẽ của công nghệ 4.0, Artificial Intelligence (trí tuệ nhân tạo) hay cụ thể là Machine Leaming (máy học) cũng đang từng bước phát triển không ngừng. Các ứng dụng về lĩnh vực này ngày càng được áp dụng nhiều vào đời sống con người như: Robot giúp việc, hệ thống dịch thuật, chatbox, phân loại sản phẩm...
* Mạng Neron học sâu (Deep learning Network) là lĩnh vực nghiên cứu các thuật toán, chương trình máy tính để máy tính có thể học tập và đưa ra những dự đoán như con người. Nó được ứng dụng vào nhiều ứng dụng khác nhau như khoa học, kỹ thuật, các lĩnh vực đời sống khác cũng như các ứng dụng về phân loại và phát hiện đối tượng. Một ví dụ điển hình là CNN (Convolutional Neural Network) áp dụng để nhận dụng tự động, tìm hiểu các mẫu phân biệt từ ảnh bằng cách xếp chống liên tiếp các lớp lên nhau và trong nhiều ứng dụng, CNN hiện nay được coi là trình phân loại ảnh mạnh và thúc đẩy các công nghệ trong lĩnh vực thị giác máy tính. làm đòn bảy cho quá trình học máy, cụ thể là được ứng dụng vào đề tài "Phân loại biển báo giao thông”

1. **Mô tả bài toán cần giải quyết**

* Bài toán cần giải quyết của nhận diện biển báo giao thông là xác định biển báo và phân loại biển báo trong hình ảnh hoặc video. Trong thực tế, các hình ảnh có thể chứa nhiều biển báo khác nhau với các đặc điểm riêng biệt, ví dụ như màu sắc, hình dạng và kích thước khác nhau. Do đó, bài toán nhận diện biển báo là một bài toán phức tạp và thách thức trong lĩnh vực trí tuệ nhân tạo. Việc giải quyết bài toán nhận diện biển báo có thể giúp chúng ta hiểu và tự động hóa các tác vụ như tự động phát hiện các biển báo giao thông trong hình ảnh hoặc video, hỗ trợ trong điều khiển robot tự động hoặc tự động lái xe, hoặc giúp phát triển các ứng dụng tương tác với người dùng. Để giải quyết bài toán nhận diện biển báo, các phương pháp và mô hình máy học được sử dụng để huấn luyện các mô hình nhận diện biển báo, đào tạo các mô hình này trên các tập dữ liệu lớn để học cách phát hiện đặc điểm của các đối tượng khác nhau và phân loại chúng vào các lớp tương ứng.
* Để thực hiện nhận diện biển báo, các bài toán cần được giải quyết bao gồm:

1. Xây dựng một bộ dữ liệu đủ lớn và đa dạng: Bộ dữ liệu cần phải bao gồm nhiều hình ảnh hoặc video của các biển báo khác nhau, với nhiều góc chụp, ánh sáng và môi trường khác nhau. Có tất cả 43 classes.
2. Xác định các đặc trưng quan trọng: Các đặc trưng quan trọng của đối tượng cần được xác định và trích xuất ra từ ảnh hoặc video, chẳng hạn như hình dạng, màu sắc, vị trí, kích thước, texture, ...
3. Xây dựng mô hình học máy: Mô hình học máy được xây dựng dựa trên các đặc trưng quan trọng được trích xuất. Ở đây em sử dụng mô hình máy học sâu mạng neural và các mô hình truyên thống khác.
4. Huấn luyện mô hình: Mô hình được huấn luyện trên bộ dữ liệu đã chuẩn bị. Quá trình huấn luyện được thực hiện thông qua các phương pháp như backpropagation hoặc gradient descent.
5. Đánh giá và điều chỉnh mô hình: Mô hình được đánh giá dựa trên các chỉ số đánh giá như độ chính xác, độ nhạy, độ đặc hiệu, và điều chỉnh nếu cần thiết để tăng độ chính xác.
6. Sử dụng mô hình: Sau khi huấn luyện và đánh giá mô hình, nó có thể được sử dụng để nhận diện biển báo trong các ảnh hoặc video mới.
7. **Dữ liệu và nguồn dữ liệu**

* Mô hình đã được huấn luyện để nhận diện đối tượng được lưu trong tệp keras\_model.h5. Khi chạy chương trình, mô hình được tải vào bằng hàm load\_model từ thư viện Keras.
* Tệp H5 là tệp dữ liệu được lưu ở Định dạng dữ liệu phân cấp (HDF). Nó chứa các mảng dữ liệu khoa học đa chiều. Các tệp H5 thường được sử dụng trong hàng không vũ trụ, vật lý, kỹ thuật, tài chính, nghiên cứu học thuật, bộ gen, thiên văn học, dụng cụ điện tử và lĩnh vực y tế..
* Nhãn của các đối tượng được lưu trong tệp labels.txt. Tệp này chứa danh sách các tên đối tượng mà mô hình đã được huấn luyện để nhận diện.
* Ảnh hoặc video từ camera được sử dụng để thực hiện nhận diện đối tượng. Trong đoạn code, ảnh từ camera được đọc bằng OpenCV và được chuyển đổi thành định dạng Image TK để hiển thị trên giao diện người dùng.
* Các thư viện Tkinter, threading, numpy, PIL cũng được sử dụng để hiển thị giao diện người dùng và xử lý ảnh.
* Nguồn dữ liệu là ảnh hoặc video đọc từ camera được kết nối với máy

1. **Các đặc trưng được sử dụng**

* Đặc trưng được sử dụng trong nhận diện đồ vật phụ thuộc vào cách thức huấn luyện mô hình và loại dữ liệu được sử dụng để huấn luyện. Trong mô hình Deep Learning được sử dụng trong đoạn code, đặc trưng được sử dụng là đặc trưng hình ảnh. Đặc trưng này được trích xuất thông qua việc sử dụng mạng Convolutional Neural Network (CNN) và được sử dụng để phân loại đối tượng. Trong quá trình huấn luyện mô hình, mạng CNN được cung cấp dữ liệu hình ảnh và nhãn tương ứng của đối tượng trong hình ảnh đó. Mô hình sẽ học cách trích xuất các đặc trưng phổ biến từ hình ảnh và phân loại chúng thành các lớp khác nhau. Khi áp dụng mô hình cho ảnh mới, mô hình sử dụng đặc trưng của ảnh đó để dự đoán đối tượng trong ảnh.
* **CNN (Convolutional Neural Network)** là một loại mạng nơ-ron truyền thống thường được sử dụng trong các bài toán liên quan đến xử lý hình ảnh và video. CNN được thiết kế để tự động học và trích xuất các đặc trưng (features) từ ảnh thông qua việc sử dụng các lớp tích chập (convolutional layers) để tìm kiếm các đặc trưng tương ứng với các kích thước khác nhau của ảnh.
* Mỗi lớp tích chập trong CNN được thiết kế để nhận vào các tín hiệu ảnh và trích xuất các đặc trưng tương ứng với một số kernel (nhân tích chập) được định trước. Sau khi qua các lớp tích chập, kết quả được thu được là các feature map, các bản đồ đặc trưng mà mỗi bản đồ đó biểu diễn cho một loại đặc trưng cụ thể của ảnh.
* Sau đó, các feature map sẽ được đưa qua các lớp pooling để giảm chiều dữ liệu và làm giảm độ phức tạp tính toán. Cuối cùng, dữ liệu được đưa qua các lớp fully connected để phân loại.
* Tkinter là thư viện GUI tiêu chuẩn cho Python. Tkinter trong Python cung cấp một cách nhanh chóng và dễ dàng để tạo các ứng dụng GUI. Tkinter cung cấp giao diện hướng đối tượng cho bộ công cụ Tk GUI.
* OpenCV là tên viết tắt của open source computer vision library – có thể được hiểu là một thư viện nguồn mở cho máy tính. Cụ thể hơn OpenCV là kho lưu trữ các mã nguồn mở được dùng để xử lý hình ảnh, phát triển các ứng dụng đồ họa trong thời gian thực. OpenCV cho phép cải thiện tốc độ của CPU khi thực hiện các hoạt động real time. Nó còn cung cấp một số lượng lớn các mã xử lý phục vụ cho quy trình của thị giác máy tính hay các learning machine khác.
* PIL: Python Imaging Library là một thư viện Python dùng để xử lý ảnh.
* Keras: là một thư viện mã nguồn mở sử dụng để xây dựng và huấn luyện mô hình Deep Learning.
* Threading: Thư viện Python được sử dụng để lập trình đa luồng (multi-threading), cho phép chương trình thực hiện nhiều tác vụ đồng thời.
* Numpy: Thư viện Python sử dụng để thao tác với mảng số, ma trận và các phép tính toán trên đó.

1. **Phương pháp đánh giá**

* Trong Machine Learning, phương pháp đánh giá được sử dụng để đánh giá hiệu suất của một mô hình dự đoán trên tập dữ liệu chưa được nhìn thấy trước đó. Một mô hình được đào tạo trên tập dữ liệu đã biết, vì vậy phương pháp đánh giá được sử dụng để kiểm tra khả năng của mô hình dự đoán trên tập dữ liệu chưa được sử dụng để đào tạo.
* Một số phương pháp đánh giá thông thường bao gồm:

1. Chia tập dữ liệu: Tách tập dữ liệu thành tập huấn luyện và tập kiểm tra. Mô hình được huấn luyện trên tập huấn luyện và đánh giá trên tập kiểm tra để đánh giá hiệu suất của nó trên dữ liệu chưa được sử dụng.
2. Cross-validation: Một phương pháp đánh giá mô hình bằng cách chia tập dữ liệu thành nhiều tập con (folds), mô hình được huấn luyện trên tập dữ liệu bao gồm nhiều folds và đánh giá trên tập fold còn lại để đánh giá hiệu suất trên dữ liệu chưa được sử dụng. Phương pháp này có thể giúp đánh giá hiệu suất của mô hình một cách chính xác hơn.
3. K-fold cross-validation: Một dạng đặc biệt của phương pháp cross-validation, trong đó tập dữ liệu được chia thành k folds và mô hình được huấn luyện và đánh giá trên k lần với các fold khác nhau.
4. Confusion matrix: Là một ma trận được sử dụng để đánh giá hiệu suất của mô hình phân loại. Ma trận này hiển thị số lượng các mẫu dữ liệu được phân loại đúng và sai, dựa trên đầu ra được dự đoán của mô hình và đầu ra thực tế của dữ liệu.
5. ROC curve: Là một đường cong được sử dụng để đánh giá hiệu suất của mô hình phân loại dựa trên độ nhạy và đặc hiệu của mô hình. ROC curve biểu diễn tỷ lệ giữa True Positive Rate (TPR) và False Positive Rate (FPR) với các ngưỡng quyết định khác nhau.
6. **Cách bố trí thực nghiệm**

* Thu thập và xử lý dữ liệu: Cần phù hợp và đảm bảo tính nhất quán và chính xác.
* Huấn luyện thuật toán: Cho máy sử dụng mô hình như mạng nơ-ron.
* Kiểm tra và đánh giá: Xác đinh hiệu quả và đánh giá khả năng.
* Tối ưu hóa thuật toán: Bằng cách điều chỉnh các thông số và cấu trúc mô hình.
* Triển khai và ứng dụng: Trên nhiều thiết bị.

1. **Phương pháp học**

* Mô hình sử dụng mô hình CNN được huấn luyện sẵn để phân loại các đối tượng trong ảnh từ camera, dữ liệu được lưu trữ trong file “.h5” và nhãn các đối tượng được lưu trong file “.txt”. Đối với mỗi khung hình từ camera mô hình tiến hành tiến hành chuẩn bị đầu vào bằng cách chuyển đổi kích thước và định dạng ảnh, sau đó truyền ảnh đó vào mô hình để dự đoán lớp của đối tượng.
* Tuy nhiên, để đánh giá chính xác hiệu suất của đoạn code, cần thực hiện kiểm tra với nhiều trường hợp khác nhau. Kết quả thực nghiệm có thể khác nhau tùy thuộc vào điều kiện môi trường, chất lượng ảnh, độ phân giải, tốc độ xử lý và độ chính xác của mô hình đã được huấn luyện.

**Phần 2: ỨNG DỤNG**

1. **Xác định yêu cầu**

* nhận diện biển báo được sử dụng để giải quyết nhiều vấn đề thực tế khác nhau như:
* Tự động phân loại biển báo: trong các cửa hàng trực tuyến hoặc siêu thị, việc phân loại sản phẩm dựa trên hình ảnh giúp tăng tốc độ và độ chính xác của quá trình phân loại hàng hoá.
* Điều khiển tự động: nhận diện biển báo cũng được sử dụng trong các hệ thống tự động hóa như xe tự lái, robot, các hệ thống cảnh báo trên các phương tiện giao thông công cộng, và nhiều ứng dụng khác.
* Trong mỗi trường hợp sử dụng, bài toán nhận diện biển báo giao thông đều có mục đích khác nhau tùy vào yêu cầu cụ thể của ứng dụng.
* Làm việc với dữ liệu từ máy ảnh, webcam, video
* Thông thường, thời gian đáp ứng yêu cầu của hệ thống nhận diện biển báo sẽ phụ thuộc vào yêu cầu cụ thể của ứng dụng. Thời gian đáp ứng có thể được đo bằng cách tính thời gian giữa khi một biển báo xuất hiện trong hình ảnh và khi hệ thống trả về kết quả nhận diện.
* Thời gian đáp ứng yêu cầu của hệ thống cần đạt được có thể khác nhau tùy thuộc vào mục đích sử dụng của hệ thống. Ví dụ, nếu hệ thống được sử dụng để giám sát an ninh, thì thời gian đáp ứng cần đạt được có thể rất nhanh, trong khi đối với các ứng dụng như chẩn đoán y tế, thời gian đáp ứng có thể chậm hơn.
* Tuy nhiên, để đảm bảo tính chính xác và độ tin cậy của kết quả, thời gian đáp ứng của hệ thống cần phải nhanh và đáp ứng được yêu cầu của ứng dụng trong thời gian thực.
* Độ chính xác yêu cầu của hệ thống nhận diện đối tượng sẽ phụ thuộc vào ứng dụng cụ thể. Ví dụ như trong các ứng dụng y tế hoặc an ninh, độ chính xác yêu cầu phải cao, tối thiểu là 95% hoặc thậm chí là 99%, trong khi đối với các ứng dụng khác như giám sát giao thông, độ chính xác có thể được yêu cầu thấp hơn. Do đó, độ chính xác yêu cầu của hệ thống cần đạt được sẽ phụ thuộc vào yêu cầu của ứng dụng cụ thể.

1. **Các chức năng**

* Chọn ảnh đề nhận diện mệnh giá.
* Sử dụng camera để nhận diện mệnh giá.

1. **Giao diện**

* Giao diện input.

Ảnh có chứa văn bản

Mô tả được tạo tự động

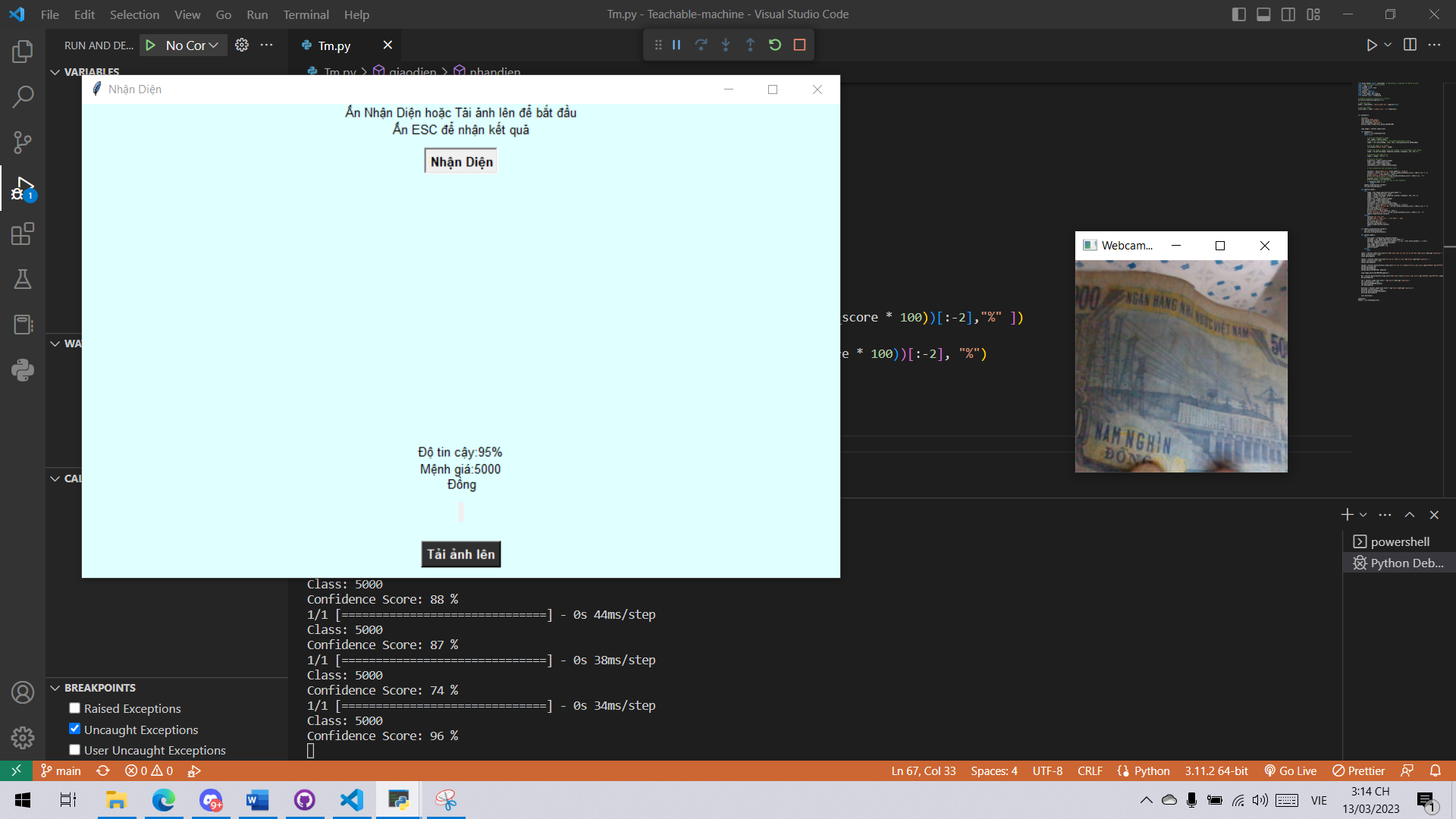
* Giao diện output.



Ảnh có chứa văn bản

Mô tả được tạo tự động

* Giao diện camera.



**Phần 3: KẾT LUẬN**

1. Ưu điểm

* Tăng năng suất và hiệu quả:
* Tiết kiêm chi phí: Giúp giảm thiểu sự phụ thuộc và lao động và giảm chi phí nhân lực và giúp giảm sai sót.
* Độ chính xác cao: Độ chính xác ngày càng được cải thiện, đảm bảo tính chính xác cao.
* Tính tự động và độc lập: Hoạt động một các độc lập, không cần sự can thiệp của con người. Giúp cho việc quản lý trở lên dễ dàng hơn.
* Tính linh hoạt: Có thể sử dụng ở nhiều lĩnh vực khác nhau.

1. Nhược điểm

* Tỉ lệ chính xác chưa cao,
* Giao diện còn đơn giản
* Dữ liệu từ mô hình train sẵn